EXHIBIT A

1 KENNETH B. WILSON (SBN 130009)
   ken@coastsidelegal.com
2 COASTSIDE LEGAL
   455 1st Avenue
3 Half Moon Bay, California 94019
   Telephone: (650) 440-4211
4
   JOSHUA M. MASUR, (SBN 203510)
5 *jmasur@zuberlaw.com*
   ZUBER LAWLER & DEL DUCA LLP
6 2000 Broadway Street, Suite 154
   Redwood City, California 94063
7 Telephone: (650) 434-8538
   Facsimile: (213) 596-5621
8
   ZACHARY S. DAVIDSON (SBN 287041)
9 *zdavidson@zuberlawler.com*
   ZUBER LAWLER & DEL DUCA LLP
10 350 S. Grand Ave., 32nd Fl.
   Los Angeles, California 90071
11 Telephone: (213) 596-5620
   Facsimile: (213) 596-5621
12

13 Attorneys for Defendant
   REDBUBBLE INC.
14

15   **UNITED STATES DISTRICT COURT**

16   **FOR THE NORTHERN DISTRICT OF CALIFORNIA**

17   **SAN FRANCISCO DIVISION**

18

19 ATARI INTERACTIVE, INC., | Case No. 3:18-cv-03451-JST

20   Plaintiff,

21   v.   **EXPERT REPORT OF**
                                          MICHAEL MASNICK
22 REDBUBBLE INC.,

23   Defendant.

24

25

26 **I.   Introduction**

27   1.   I have been asked to opine on issues facing websites like Redbubble that moderate

28 content at scale.

2.   I am a journalist, researcher, writer, and consultant who has studied and written

about issues related to online content moderation for over a decade. I have regularly written about

and speak on these issues, and have written, edited, and published papers related to the topic. I

have also conducted numerous interviews with content moderation experts, on the policy, legal

and operations teams, from a variety of different internet platforms.

3.   I am a frequent speaker on issues related to content moderation and, in particular,

the challenges of content moderation with regards to intellectual property. I have spoken before a

Copyright Office Roundtable discussion concerning an upcoming report considering changes to

Section 512 of the DMCA, exploring how well the laws currently deal with issues around

moderation and takedowns (https://www.copyright.gov/policy/section512/public-

roundtable/transcript_05-12-2016.pdf May 12, 2016). I have presented before the EU Parliament

on issues related to copyright protection and internet platforms. I have spoken regularly about

content moderation and IP issues at conferences including the Content Moderation at Scale

Summit, CES (the Consumer Electronics Show), State of the Net, Leadership Music Digital

Summit, NARM (National Association of Recording Merchandisers) Summit (now called the

Music Business Association), MIDEM (the world's largest music trade show) and many others.

Earlier this year, I gave the keynote address at the California Lawyers Association "IP and the

Internet" Conference, in which my talk was about the vast challenges for internet platforms being

required to proactively moderate copyright-protected content

(https://mk0californiala1yag7.kinstacdn.com/wp-content/uploads/2019/05/2019-IP-and-Internet-

Conference-Brochure.pdf June 6, 2019).

4.   As should be clear from the opinions I express herein, my opinions were not

developed solely, or even largely, in the context of the particular materials I reviewed for purposes

of this case. I have a longstanding interest and expertise in content moderation issues. Indeed, I

have written so frequently about the impossibility of consistently performing content moderation

at scale that the site I founded and run, Techdirt, has its own "tag" for posts on the issue. I have

reproduced, summarized, or referenced many of my postings with that tag –

https://www.techdirt.com/blog/?tag=content+moderation+at+scale – in this report. Where I have

MASNICK EXPERT REPORT

reproduced, summarized, or referenced those postings, I have done so for convenience and to keep the length of this report manageable. To the extent necessary, I incorporate the referenced postings as though they were fully set forth herein (which I understand is a legal phrase). In addition to those postings identified within the opinions themselves, the facts or data I specifically considered in forming my opinions related to this case include:

- Atari's complaint against Redbubble and Redbubble's Second Amended Answer;
- Portions of Redbubble's summary judgment briefing and factual declarations of James Toy and Ayaire Cantil-Vorhees in Redbubble's cases against Ohio State University and LTTB, LLC and of Anuj Luthra in the Ohio State case;
- The "one row per work" document that was Exhibit A to Redbubble's June 24, 2019 interrogatory responses;
- The 2011 notice that Atari sent to Redbubble (RBAT033206);
- Counternotices that Redbubble received from artists whose listings it took down (RBAT027021, RBAT027060, RBAT027345, RBAT027347);
- Redbubble's proactive policing guidelines for Atari (RBAT029889 and RBAT031848);
- Redbubble's public-facing documents regarding its policies and procedures, including
    - "Intellectual Property FAQ" (RBAT121867);
    - "Community and Content Guidelines" (RBAT124431);
    - "Content & Suspension" (RBAT124440);
    - "Copyright" (RBAT124445);
    - "Fair Use & Other Defenses" (RBAT124450);
    - "My Work Was Removed" (RBAT124457);
    - "Restricted & Deleted Accounts" (RBAT124466);
    - "Someone Is Infringing My Rights or Another" (RBAT124470);
    - "Trademark" (RBAT124474);
    - "Redbubble IP-Publicity Rights Policy" (RBAT124477);

MASNICK EXPERT REPORT

1   • "User Agreement" (RBAT124483); and

2   • "Redbubble User Policies" (RBAT124499); and

3   • Confidential documents containing Redbubble's policies and procedures for

4   content moderation, including:

5   • "Correspondence - Questions from users and rights holders"

6   (RBAT029831);

7   • "Deliberate Misuse" (RBAT029832);

8   • "Deliberate Misuse Examples" (RBAT029833);

9   • "Redbubble Policies, some Foundational Ideas, Intellectual);

10   • "Property and" (RBAT029837);

11   • "Repeat Infringer Policy" (RBAT029839);

12   • "Suspend-from-sale" (RBAT029842);

13   • "Warnings and Account Restriction" (RBAT029843);

14   • "MPI Scripts" (RBAT029845);

15   • "Scripts - Warn - Restrict - Deliberate Misuse" (RBAT029878);

16   • "Marketplace Integrity Team" (RBAT029887);

17   • "Script - response to counternotice (Atari)" (RBAT121859);

18   • "Title & Trademark Policing Tips" (RBAT121870);

19   • "Proactive Policing" (RBAT121873);

20   • "Understanding Policing Guidelines" (RBAT121880);

21   • "Moderating Content - Takedown" (RBAT121884);

22   • "Marketplace Integrity Onboarding" (RBAT121886);

23   • "Reviewing Accounts- Escalation and Account Abuse" (RBAT124546);

24   and

25   • "ZD [ZenDesk] Takedown Form Requirements" (RBAT125968).

26   5.   I understand that the applicable rules require me to identify "all publications

27   authored in the previous 10 years." Most of my writing is in the form of short pieces, including

28   blog entries, that are more informal and more numerous than what is considered a "publication" in

1 academic circles – sometimes several in a day. I estimate that I have written thousands of posts in

2 the past 10 years. I have attempted here – both in this section and in my substantive discussion

3 below – to point to the locations where lists of those posts can be found, and to identify certain

4 writings that are particularly relevant to this litigation.

5   6.   Techdirt automatically lists my publications on the site – which are too numerous

6 to list here – at the URL https://www.techdirt.com/user/mmasnick.

7   7.   I also founded and run The Copia Institute, a think tank that does research and

8 events, often with a focus on issues related to content moderation and/or intellectual property. I

9 have authored or co-authored numerous Copia papers, listed at https://copia.is/library/. My Copia

10 papers on related topics include "Don't Shoot the Message Board" (https://copia.is/library/dont-

11 shoot-the-message-board/, discussing the impact of intermediary liability protections on

12 investment), "The Sky is Rising" (https://copia.is/library/the-sky-is-rising-2019/, on the state of

13 the industry in responding to and adapting to challenges with infringement) and "The Carrot and

14 the Stick" (https://copia.is/library/the-carrot-or-the-stick/, exploring the impact of licensed internet

15 platforms on IP infringement).

16   8.   I recently published a paper for the Knight First Amendment Institute at Columbia

17 University (https://knightcolumbia.org/content/protocols-not-platforms-a-technological-approach-

18 to-free-speech, August 21, 2019) exploring technological approaches to rethinking how internet

19 platforms could be structured, including a detailed discussion on the challenges of content

20 moderation and how a different technological approach might impact those challenges.

21   9.   In addition to the above, I have occasionally been published by other news

22 publications over the past decade. These include Wired Magazine

23 (https://www.wired.com/author/mike-masnick/), The Verge

24 (https://www.theverge.com/2017/4/3/15161522/mpaa-riaa-copyright-office-library-of-congress-

25 dmca-infringement), The Hill (https://thehill.com/blogs/congress-blog/politics/315032-internet-

26 freedom-isnt-free-five-years-after-the-sopa-pipa and https://thehill.com/blogs/congress-

27 blog/technology/258258-the-answer-to-the-copyright-question-is-innovation), Reason Magazine

28 (https://reason.com/2019/12/17/can-jack-dorsey-reinvent-the-internet-by-making-twitter-more-

MASNICK EXPERT REPORT

1 like-email/), Boing Boing (https://boingboing.net/author/mikemasnick), and Morning Consult

2 (https://morningconsult.com/opinions/renegotiating-nafta-must-protect-creative-innovators/). In

3 addition to the articles above, I have contributed chapters to the book "Hacking Politics"

4 (https://www.goodreads.com/book/show/17876446-hacking-politics) and "Working Futures"

5 (https://www.goodreads.com/book/show/48342389-working-futures) which I also helped to edit.

6   10.   For the Copia Institute, I have recently been working on creating two "training

7 simulations" that both relate to content moderation. The first is a simulation for a small group of

8 individuals to have them experience what it is like to run or work on a "Trust & Safety" team in

9 charge of moderating content. This simulation is a much more advanced and detailed expansion

10 of a simplified simulation -- called "You Make the Call" that I helped put together and run at the

11 Content Moderation at Scale conference in Washington DC in 2018. I also recently helped put

12 together and run a large group simulation for the Mozilla Foundation, training content moderation

13 experts and others in what forms of disinformation they may need to deal with during the 2020 US

14 Presidential election.

15   11.   During the previous 4 years, I have testified as an expert at trial in United States v.

16 Wexler (8:15-cr-00122) in the Central District of California.

17   12.   I am being compensated at a rate of $300 per hour for my work in connection with

18 this case.

19

20 **II.   Content Moderation**

21   13.   The process by which Internet companies that host user-submitted content decide

22 which of that content to allow to be accessed and which to remove is generally called "content

23 moderation."

24   14.   Failing to moderate enough can lead to or promote major social problems, like

25 dissemination of false information, harassment, abuse or unwanted commercial messages

26 (colloquially referred to as "spam"). But moderating too much can suppress socially beneficial

27 content, like negative but true consumer reviews, criticism, commentary, parody and more.

28

1   15.   Part of the challenge is that no one agrees what is the "correct" level of moderation.

2 Ask 100 people and you will likely get 100 different answers (which I confirmed empirically in

3 2018 by getting approximately 100 content moderation experts, attending the "Content

4 Moderation at Scale" conference, to evaluate content on a lightly fictionalized social media

5 platform). What many people think must be mostly "black and white" choices actually have a

6 tremendous amount of gray. Even if there **were** clear and easy choices to make, (which in many

7 situations is not the case) at the scale of most major platforms, even a **tiny** error rate (of either

8 false positives or false negatives) will still be a very large absolute number of "mistakes."

9   16.   Even when there are clear choices to make content moderation suffers from two

10 types of errors. Using standard statistical terminology, blocking content that shouldn't be blocked

11 is called a "type I error," and failing to block content that probably should be blocked is referred to

12 as a "type II error."

13   17.   As difficult as it is to perform content moderation at a small scale, scaling content

14 moderation – doing so consistently for thousands of items of user-generated content – is

15 impossible to do well. Among other issues, both type I and type II errors are essentially

16 inevitable, and the need to involve multiple human beings, with different interpretations of often

17 subjective rules and policies, means that inconsistency is unavoidable.

18   18.   As Techdirt contributor Professor Eric Goldman of the Santa Clara University

19 School of Law has written, in a piece titled "Top Myths About Content Moderation,"

20 https://www.techdirt.com/articles/20191015/14542243198/top-myths-about-content-

21 moderation.shtml (Oct. 16, 2019), attempts to regulate, legislate, or litigate content moderation

22 often fail because of a misunderstanding of how content moderation actually works. As Prof.

23 Goldman points out, "Regulators are right to identify content moderation as a critically important

24 topic. However, until regulators overcome these myths, regulatory interventions will cause more

25 problems than they solve."

26   19.   Prof. Goldman identifies nine particular myths:

27   *Myth*: Content moderation can be done perfectly.

28

MASNICK EXPERT REPORT

Reality: Regulators routinely assume Internet services can remove all bad content without suppressing any good content. Unfortunately, they can't. First, mistakes occur when the service lacks key contextual information about the content—such as details about the author's identity, other online and offline activities, and cultural references. Second, any line-drawing exercise creates mistake-prone border cases because users routinely submit "edgy" content. Third, a high-volume service will make many mistakes, even if it's highly accurate—1 billion submissions a day at 99.9% accuracy still yields a million mistakes a day.

Myth: Bad content is easy to find and remove.

Reality: Regulators often assume every item of bad content has an impossible-to-miss flashing neon sign saying "REMOVE THIS CONTENT," but that's rare. Content is often obviously bad only in hindsight or with context unavailable to the service. Regulators' cherry-picked anecdotes don't prove otherwise.

Myth: Technologists just need to "nerd harder."

Reality: Filtering and artificial intelligence play important roles in content moderation. However, technology alone cannot magically solve the problem. "Edgy" and contextless content vexes the machines, too.

Myth: Internet services should hire more humans to review content.

Reality: Humans have biases and make mistakes too, so adding human reviewers won't lead to perfection. Furthermore, human reviewers sometimes experience an unrelenting onslaught of horrible content to protect the rest of us.

Myth: Internet companies have no incentive to moderate content.

Reality: In 1996, Congress passed 47 U.S.C. 230, which says Internet services generally aren't liable for third-party content. Due to this legal protection, critics often assume Internet services won't invest in content moderation; and some companies have stoked that perception by publicly positioning themselves as "neutral" technology platforms. Yet, virtually every Internet service moderates content, and major services like Facebook and YouTube employ many thousands of content reviewers. Why? The services have their own reputation to manage, and they care about how content can affect their users (e.g., Pinterest combats content that promotes eating disorders). Furthermore, advertisers won't let their ads appear on bad content, which provides additional financial incentives to moderate.

Myth: Content moderation, if done right, will make everyone happy.

Reality: By definition, content moderation is a zero-sum game. Someone gets their desired outcome, and someone else doesn't—and those folks won't be happy with the result.

Myth: There is a one-size-fits-all approach to content moderation.

Reality: Internet services cater to diverse audiences that have different moderation needs. For example, an online crowdsourced encyclopedia like Wikipedia, an open-source software repository like GitHub, and a payment service for content publishers like Patreon all solve different problems for their communities. These services shouldn't have identical content moderation rules.

Myth: Imposing content moderation requirements will stick it to Google and Facebook.

Reality: Google and Facebook have enough money to handle virtually any requirement imposed by regulators. Startup enterprises do not. Increased content moderation burdens are more likely to block new entrants than to punish Google and Facebook.

Myth: Poor content moderation causes anti-social behavior.

Reality: Poorly executed content moderation can accelerate bad behavior, but often the Internet simply mirrors existing anti-social behavior or tendencies. Better content moderation can't fix problems that are endemic in the human condition.

III.   Masnick's Impossibility Theorem: Content Moderation at Scale Is Impossible to Do Well

20.   On November 20, 2019, I published a piece at

https://www.techdirt.com/articles/20191111/23032743367/masnicks-impossibility-theorem-content-moderation-scale-is-impossible-to-do-well.shtml, titled "Masnick's Impossibility Theorem: Content Moderation At Scale Is Impossible To Do Well," from which this section of my report is adapted.

21.   For many years, I have explained that while many people claim that content moderation is **difficult**, that's misleading. It implies that more hard work can "solve" the mistakes made in content moderation. Instead, it is important to recognize: **content moderation at scale is impossible to do well**. Importantly, this is **not** an argument that everyone should do nothing. Nor is it an argument that companies can't do *better jobs* within their own content moderation efforts.

1   But there's a huge problem in that many people expect that these companies not only can, but

2   should, strive for a level of content moderation that is simply impossible to reach.

3   22.   I have recently proposed "Masnick's Impossibility Theorem," as a sort of play on

4   Arrow's Impossibility Theorem.[1] **Content moderation at scale is impossible to do well**. More

5   specifically, it will always end up frustrating very large segments of the population and will

6   always fail to accurately represent the "proper" level of moderation of anyone. While I have not

7   gone through the process of formalizing the theorem, as Arrow did in his seminal text, I will

8   highlight a few points on why the Impossibility Theorem regarding content moderation is

9   inevitably true.

10   23.   First, any moderation is likely to end up angering those who are moderated. Those

11   who initially posted the content in the first place almost certainly thought it properly belonged

12   wherever it was posted -- so they will likely disagree with the decision to moderate it. Some might

13   argue the obvious response to this concern is to do no moderation at all. However, that also fails

14   for the obvious reason that many people would greatly prefer some level of moderation, especially

15   given that any unmoderated area of the internet quickly fills up with spam, not to mention abusive

16   and harassing content. Or, worse, if content is violating some laws, a platform is unlikely to even

17   have the option of not moderating. Some have suggested that pushing out the moderation to the

18   ends of the network (i.e., giving more control to the end users to make their own moderation

19   choices) is better, but that also has some complications in that it puts the burden on end users, and

20   they have neither the time nor inclination to continually tweak their own settings. It also fails to

21   deal with issues regarding content that violates the law. No matter what path is chosen, it will end

22   up being not ideal for a large segment of the population. Moderation is necessary (indeed, in some

23   cases, required to stay on the right side of the law), and as such, will upset at least some significant

24   portion of those moderated.

25

26

---

27   [1] Arrow's Impossibility Theorem is a theorem put forth by Nobel Prize-winning economist Kenneth Arrow in his
doctoral dissertation and later in his 1951 book "Social Choice and Individual Values" and roughly speaking shows

28   that no voting system can accurately represent the preferences of the voters, and that no matter how a voting system
was constructed, it would end up being "unfair" to a certain set of voter preferences.

Case No. 3:18-cv-03451-JST

MASNICK EXPERT REPORT

1   24.   Second, moderation is, inherently, a subjective practice. Despite some people's

2 desire to have content moderation be more scientific and objective, true objectivity is impossible.

3 By definition, content moderation is always going to rely on judgment calls, and many of the

4 judgment calls will end up in gray areas where lots of people's opinions may differ greatly. Indeed,

5 one of the problems of content moderation that we've seen over the years is that to make good

6 decisions you often need a tremendous amount of context, and there's simply no way to adequately

7 provide that at scale in a manner that actually works. That is, when doing content moderation at

8 scale, you need to set rules, but rules leave little to no room for understanding context and

9 applying it appropriately. And thus, you get many edge cases that end up looking bad. Empirical

10 data support this conclusion. In 2018, when we turned an entire conference of "content

11 moderation" specialists into content moderators for an hour, we found that there were exactly zero

12 cases where we could get all attendees to agree on what should be done in any of the eight cases

13 we presented. Indeed, all eight scenarios involved four possible options for the content

14 moderators to choose from, and in every single case, all four options were chosen by at least some

15 participants.

16   25.   Third, people frequently underestimate the impact that "scale" has on this equation.

17 Getting 99.9% of content moderation decisions at an "acceptable" level probably works acceptably

18 for situations when you're dealing with 1,000 moderation decisions per day, but large platforms

19 are dealing with way more than that. If you assume that there are 1 million decisions made every

20 day, even with 99.9% "accuracy" (and 99.9% accuracy is already impossible), you're still going to

21 "miss" 1,000 calls. And the numbers can scale up massively from there. On Facebook alone a

22 recent report (https://www.businessinsider.com/facebook-350-million-photos-each-day-2013-

23 9?IR=T) noted that there are 350 million photos uploaded every single day. And that's just photos.

24 If there's a 99.9% accuracy rate, it's still going to make "mistakes" on 350,000 images each and

25 every day. Those add up.

26   26.   Even if you could achieve such high "accuracy," (which is unlikely) with so many

27 mistakes in absolute terms, it is always possible for someone, such as a journalist, to go searching

28 and find a bunch of those mistakes -- and point them out. This may make some suggest that if a

MASNICK EXPERT REPORT

1   reporter can find those bad calls, Facebook should be able to as well. That leaves out that

2   Facebook DID find that other 99.9%. Obviously, these numbers are just illustrative, but the point

3   stands that when you're doing content moderation at scale, the scale part means that even if you're

4   extraordinarily good, you will still make a large number of mistakes in absolute numbers every

5   single day.

6   27.   It is important to keep exploring different approaches to content moderation, and to

7   call out failures when they (frequently) occur, but it's important to recognize that there is no

8   perfect solution to content moderation, and any company, no matter how thoughtful and deliberate

9   and careful is going to make many mistakes.

10   28.   I have written repeatedly about failed content moderation at scale. Often,

11 intentionally bad actors can game content moderation policies to their own benefit, at the expense

12 of those who do not seek to manipulate those policies to their benefit. Here are just a few

13 examples from just the past few months:

14   •   White supremacists were accused of manipulating Twitter's content moderation

15   policies to get people they dislike banned from Twitter over old, out-of-context

16   jokes that were on the site for many years:

17   (https://www.techdirt.com/articles/20191201/00375943479/content-moderation-

18   scale-is-impossible-that-time-twitter-nazis-got-reporter-barred-twitter-over-some-

19   jokes.shtml Dec. 4, 2019);

20   •   Facebook, while invoking a policy to ban "anti-vaccine" advertisements, instead

21   removed pro-vaccine ads from hospitals and other health organizations, while

22   allowing ads by anti-vaccine conspiracy theorists to continue

23   (https://www.techdirt.com/articles/20191026/23243243268/content-moderation-

24   scale-remains-impossible-vaccines-edition.shtml Oct. 31, 2019); and

25   •   Men are gaming dating app Tinder's moderation policies to get women who reject

26   them banned from Tinder by falsely reporting that they have engaged in prohibited

27   behavior on the platform.

28   (https://www.techdirt.com/articles/20190918/00363543015/jerks-reporting-

12

1   women-who-swipe-left-them-tinder-once-again-highlighting-how-content-

2   moderation-gets-abused.shtml Sep. 25, 2019).

3   29.   Even without intentional abuse of reporting mechanisms, content moderation

4   policies can create perverse, or at least undesired, effects.   Famous art is removed as violating

5   policies against nudity and other explicit content

6   (https://www.techdirt.com/articles/20190818/17594042811/content-moderation-is-impossible-

7   facebook-settles-legal-fight-over-famous-painting-womans-genitals.shtml Sep. 13, 2019).

8   Because the exact same content can mean very different things in different contexts, policies

9   intended to exclude hateful content end up excluding those who seek to document or teach about

10   that hate (https://www.techdirt.com/articles/20190606/14122842345/impossibility-content-

11   moderation-youtubes-new-ban-nazis-hits-reporter-who-documents-extremism-professor-teaching-

12   about-hitler.shtml Jun. 7, 2019), or to satirize or parody it

13   (https://www.techdirt.com/articles/20190613/02505742390/content-moderation-is-impossible-

14   you-cant-expect-moderators-to-understand-satire-irony.shtml Jun 13, 2019).   In the name of

15   "stopping" terrorist content online, Facebook has deleted evidence of war crimes

16   (https://www.techdirt.com/articles/20190513/17005042202/flip-side-to-stopping-terrorist-content-

17   online-facebook-is-deleting-evidence-war-crimes.shtml May 20, 2019).   Talking about racism gets

18   blocked as hate speech (https://www.techdirt.com/articles/20190426/00092142092/impossible-

19   content-moderation-dilemmas-talking-about-racism-blocked-as-hate-speech.shtml May 1, 2019).

20   Put another way, algorithms can't tell the difference between "bad stuff" and "reporting about bad

21   stuff." (https://www.techdirt.com/articles/20180321/00413939466/once-again-algorithms-cant-

22   tell-difference-between-bad-stuff-reporting-about-bad-stuff.shtml Mar. 28, 2018).   And in an act

23   of almost unimaginable irony, Google has refused to allow AdSense words to appear with an

24   article we published about the "impossible choices" platforms have to make when moderating

25   speech on their platforms. (https://www.techdirt.com/articles/20190108/00410741352/google-still-

26   says-our-post-content-moderation-is-dangerous-derogatory.shtml Jan. 10, 2019;

27   https://www.techdirt.com/articles/20181024/09444240903/google-says-our-article-difficulty-

28   good-content-moderation-is-dangerous.shtml Oct. 24, 2018;

1   https://www.techdirt.com/articles/20180808/17090940397/platforms-speech-truth-policy-

2   policing-impossible-choices.shtml Aug. 8, 2018). There may be no better example than

3   "moderating" an article about how internet platforms to illustrate how bad most internet platforms

4   are at content moderation.

5   30.   Creating categories of unmoderated content creates problems of its own. Elizabeth

6   Warren's presidential campaign has documented that Facebook's refusal to moderate candidate

7   political ads fosters abuse (https://www.techdirt.com/articles/20191014/22010943192/elizabeth-

8   warrens-feud-with-facebook-over-false-ads-just-highlights-impossibility-content-moderation-

9   scale.shtml Oct. 15, 2019). This happened a few months after Elizabeth Warren's presidential

10   campaign complained about her own campaign's ads being moderated

11   (https://www.techdirt.com/articles/20190312/00332641776/everyones-overreacting-to-wrong-

12   thing-about-facebook-briefly-blocking-elizabeth-warrens-ads.shtml March 12, 2019). Those two

13   stories alone help demonstrated the impossibility of content moderation. Warren's campaign was

14   upset when Facebook moderated her ads, saying that the company had "too much power" and

15   shouldn't be deciding what political ads are allowed. Yet when Facebook chose not to moderate

16   other advertisements (those making misleading claims about her), she got upset that the company

17   chose not to act.

18   31.   Implementation of any content moderation policy is not nearly as easy as many

19   people believe. Almost every content moderation decision involves some sort of subjective call.

20   Is this content promoting bad behavior or documenting bad behavior? Is this content making a

21   joke at someone's expense, or is it a part of friendly in-group banter? Is this content infringing, or

22   is it protected fair use? The end result of all of these subjective decisions is that "mistakes" are

23   made. Archival and documentary footage documenting bad behavior gets blocked as promoting it.

24   Joking among friends gets blocked as an attack. Protected parodies get blocked as infringement.

25   All of those mistakes can have longer term implications on free speech, education, and culture.

26   32.   These impossible decisions have been around for many years. Early attempts to

27   ban "porn" also took down information on breast cancer. Attempts to block "terrorist content"

28   have repeatedly taken down human rights activists documenting war crimes. Internet content

1   moderation has always had problems, and for each "fix" to a difficult scenario, only more

2   problems are introduced.

3   33.   A report by the podcast Radiolab, talking to Facebook moderators

4   (https://www.wnycstudios.org/podcasts/radiolab/articles/post-no-evil August 17, 2018 and

5   covered on Techdirt https://www.techdirt.com/articles/20180817/11460840450/before-you-talk-

6   about-how-easy-content-moderation-is-you-should-listen-to-this.shtml August 21, 2018)

7   highlights just a small set of examples, beginning with Facebook's initial "no nudity" policy.

8   34.   The team at Facebook then realized they had to define "nudity," which an early

9   Facebook employees described as the following: "So, I mean, yeah, first cut at it was visible male

10   and female genitalia and then visible female breasts. And then the question is, well, okay, how

11   much of her breast needs to be showing before it's nude? And the thing that we landed on was if

12   you could see essentially the nipple and areola, then that's nudity." But that created problems for

13   breastfeeding mothers who felt their images were unfairly blocked.

14   35.   Facebook then adjusted the rules to have what was called an "attachment clause"

15   for breastfeeding. If the baby was "attached," then the image could stay up. But that upset

16   mothers who had posted pictures of them with their babies right after they'd finished

17   breastfeeding. So Facebook had to add another rule for that situation. But then they found that

18   people started posting pornographic pictures that looked the same as baby breastfeeding images,

19   but with adult males. So Facebook then decided to put an age cap on the person breastfeeding --

20   but they chose "toddler" as the age cap, which is already subjective and actually below the World

21   Health Organization's recommendations for the proper age to stop breastfeeding.

22   36.   The story becomes even more absurd, and more representative of the impossibility

23   of moderating content at scale, when a Facebook employee comes across an image of a teenage

24   girl "African by dress and skin, breastfeeding a baby goat." It turns out that this is a cultural,

25   traditional survival practice in Kenya for surviving droughts. But for Facebook, moderators had no

26   idea how it fit with their increasingly long list of rules regarding breastfeeding -- leading them to

27   have change the rules yet again to deal with this new, unexpected variation, that all previous rules

28   and amendments failed to contend with.

1   37.   As Facebook and most other platforms eventually discover, every content

2 moderation rule leads to new questions, new edge cases, and new variations -- many of which

3 require subclauses, asterisks and amendments.

4   38.   Indeed, the Supreme Court itself seemed to recognize this in defining obscenity

5 with Justice Potter Stewart's now iconic phrase "I know it when I see it" in Jacobellis v. Ohio.

6 That was, in effect, a Supreme Court Justice admitting that he could not come up with a written

7 policy or rules to define obscenity, and that it was always going to come down to the subjective

8 opinion of whoever was judging that content.

9   39.   Yet in this world, where we expect every internet platform to moderate all of the

10 content on their platform, "I know it when I see it" is not sufficient. But any other set of rules, no

11 matter how detailed and flexible, can never account for all variations or all subjective decisions.

12

13 **IV.   Teespring's Takedown of Techdirt's "Copying Is Not Theft" Products**

14   40.   As the editor of Techdirt and CEO of its parent company, Floor64. Inc., I have

15 personal experience with the consequences of overbroad moderation. Beginning in 2016, we

16 offered t-shirts and other products with a "copying is not theft"[2] graphic on our Teespring

17 storefront.

18

19

20

21

22

23

24

25

---

26   [2] By saying "copying is not theft," I and Techdirt are attempting to underscore the distinction between copying and theft. Stealing an item deprives the rightful owner of that item, while making

27 a copy of the item leaves the item from which the copy was made. The fact that copying is not theft does not necessarily make copying legally or morally right or defensible. But even when

28 copying is unlawful, it is not theft. Particularly in the digital age, this distinction is often meaningful.

16   Case No. 3:18-cv-03451-JST

MASNICK EXPERT REPORT

41.   In early December 2019, Techdirt received a notice from Teespring that our "copying is not theft" products had been taken down by Teespring for alleged copyright infringement. I've attached our takedown correspondence with Teespring as Exhibit A. We emailed Teespring's "IP Escalations" address in the takedown notification, clarifying that we had created and owned the original design and it did not include any "third-party content." But rather than reinstate our products, Teespring responded that they could not reinstate the campaign because it violates Teespring's "Acceptable Use Policies" (https://teespring.com/policies/acceptable-use) against nudity, hate speech, violence, "promot[ing] illegal activity," and "false or misleading claims," as well as intellectual property rights. When we asked which policy we violated, Teespring replied with a link to the same policy page. When we clarified that we had read the page, that we believed that we were not in violation of any of the policies, and asked which of the six policies we had supposedly violated, we received what appeared to be a scripted customer service response:

> We apologize if you disagree with our decision and for any inconvenience this matter has caused. Please understand that we are not in a position to debate our policies or discuss this issue further; however, your feedback has been noted and we truly appreciate your time today.

MASNICK EXPERT REPORT

42.   When we repeated our request that Teespring identify the policy we violated, Teespring responded:

> You've been advised three times that the content has violated our acceptable use policy. You have been provided with the links to this policy for further information. This policy and choice to remove the content are not up for discussion. We apologize if you disagree with the decision. You will not receive anymore communication from us on this matter.

43.   Due to Teespring's takedown of our products, we moved our Copying Is Not Theft store to Threadless, at https://techdirt.threadless.com/designs/copying-is-not-theft/.

V.   Excessive Moderation Is A Significant And Persistent Problem When It Comes To Copyright

44.   The situation with Teespring was hardly the first time that we experienced issues with excessive copyright moderation. Over the years, we have dealt with multiple attempts by individuals who sought to use clearly erroneous copyright takedown notices in an attempt to censor or silence articles and posts they do not like.

45.   In 2017 we reported on what appeared to be a crowdfunding scam, seeking to raise money for a transcription device, called Titan Note, that had initially raised over $1 million despite many questions about whether or not the technology actually existed. The individual behind the project sent us a DMCA takedown notice claiming that our discussion was infringing by including screenshots from his website and Facebook comments (https://www.techdirt.com/articles/20170531/16392437490/our-response-to-titan-note-sending-frivolous-takedown-notice-over-our-critical-coverage.shtml June 1, 2017).

46.   There are numerous stories of potentially abusive DMCA takedown claims, such that we have many of them tagged as "Copyright as Censorship" (https://www.techdirt.com/search-g.php?num=20&q=copyright+as+censorship&search=Search) used to describe cases where DMCA takedown or other copyright claiming systems were used not to stop the infringement of actual copyright infringement, but rather to silence commentary or criticism. I incorporate my posts under that tag as part of this report.

47.   FIFA, the organization that runs the Soccer World Cup last year used copyright claims to take down a video of protestors during a World Cup match (https://www.techdirt.com/articles/20180716/16190840245/copyright-as-censorship-fifas-overaggressive-copyright-takedowns-target-fans-celebrating-pussy-riot-protesting.shtml July 17, 2018). A reporter for the tech news website Venture Beat issued a DMCA takedown notice to Twitter to get criticism of one of his articles taken down (https://www.techdirt.com/articles/20180509/17341739812/venture-beat-reporter-abuses-dmca-to-silence-critic.shtml May 10, 2018). A video game company, Wild Game Studio issued a DMCA takedown notice to YouTube to remove a critical game review (https://www.techdirt.com/articles/20131021/00080224939/copyright-as-censorship-again-game-developer-takes-down-scathing-youtube-review.shtml October 21, 2013). A blog covering copyright law in India got targeted with an erroneous DMCA takedown notice for discussing the copyright issues around a song. The copyright holder for that song issued the takedown notice (https://www.techdirt.com/articles/20190307/17360241761/bogus-dmca-takedown-targeting-indian-copyright-blog-demonstrates-problems-notice-takedown.shtml March 14, 2019). The open document hosting platform Scribd took down the public domain Mueller Report in the false belief that someone held the copyright on it (https://www.techdirt.com/articles/20190420/23233942052/scribds-takedown-public-domain-mueller-report-is-preview-eus-future-under-copyright-directive.shtml April 22, 2019). The TV channel Starz used the DMCA to attempt to take down a news article that merely spoke about some of Starz's TV shows being leaked on the internet. The article did not include any of the copyright-covered content, nor did it provide links, details or instructions on how to find the content. It was merely reporting the news that the content had leaked and Starz sent a DMCA takedown notice to Twitter, seeking to have a tweet about the article existing removed from Twitter (https://www.techdirt.com/articles/20190414/23472142010/starz-really-really-doesnt-want-you-to-know-that-torrentfreak-wrote-about-leaked-shows-that-anyone-tweeted-about-it.shtml April 15, 2019). A convicted investment advisor sought to use the DMCA takedown process to remove the news of his own conviction from Google, asking it to delist links to the FBI,

MASNICK EXPERT REPORT

1   the DOJ, and Wikipedia that all discussed his arrest and conviction

2   (https://www.techdirt.com/articles/20180502/18304039765/another-convicted-fraudster-attempts-

3   to-manage-his-reputation-with-bogus-dmca-takedown-notices.shtml May 7, 2018).

4   48.   While these present a variety of anecdotal evidence of excessive or abusive

5   copyright takedowns -- which highlight the difficult that internet platforms have in determining

6   whether or not a takedown request is legitimate, there is also empirical data to support this. In

7   2016, researchers Jennifer Urban, Joe Karagais, and Brianna Schofield, released a study called

8   "Notice and Takedown in Everyday Practice."

9   (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2755628) Among other things, the report

10 found widespread problems with copyright takedown notices. From the report (emphasis mine):

11 - One in twenty-five of the takedown requests (4.2%) **were fundamentally flawed**

12   **because they targeted content that clearly did not match the identified infringed**

13   **work. This extrapolates to approximately 4.5 million requests** suffering from

14   this problem across the entire six-month dataset.

15 - Nearly a third of takedown requests (28.4%) **had characteristics that raised clear**

16   **questions about their validity**, based solely on the facial review and comparisons

17   we were able to conduct. Some had multiple potential issues. While these requests

18   cannot be described as categorically invalid without further investigation, **they**

19   **suggest that a very substantial number of requests in the six-month dataset—**

20   **approximately 30.1 million—would benefit from human review.**

21 - This "questionable" set included requests that raised questions about compliance

22   with the statutory requirements (15.4%), potential fair use defenses (7.3%), and

23   subject matter inappropriate for DMCA takedown (2.3%), along with a small

24   handful of other issues.

25

26 **VI.**   **There Is No Magic Bullet for Moderating A Social Media Platform**

27   49.   On May 18, 2018, I published a piece at

28 https://www.techdirt.com/articles/20180518/00271539858/there-is-no-magic-bullet-moderating-

social-media-platform.shtml, titled "There Is No Magic Bullet For Moderating A Social Media

Platform," from which this section of my report is adapted.

50.   We frequently see people who suggest that the reason social media platforms are *so*

bad at moderating content is because they just don't care or don't try hard enough. While it is true

that these platforms can often do a much better job, it's still amazing at how many people think

that deciding what content "belongs" and what content doesn't belong is somehow easy. In May

2018, in Washington DC there was the **Content Moderation at Scale** "COMO" conference

(http://comoatscale.com/). It was a one day event in which many internet platform companies

revealed (sometimes for the first time) how they handle questions around content moderation. It

was a follow-up to a similar event at Santa Clara University held back in February 2018

(http://law.scu.edu/event/content-moderation-removal-at-scale/) (for which I was the editor and

Techdirt was the official publisher of the multiple academic and industry papers that were written

in conjunction with the event: https://www.techdirt.com/articles/20180128/11121539097/time-to-

talk-about-internet-companies-content-moderation-operations.shtml January 29, 2018)

51.   For the DC event, we teamed up with the Center for Democracy and Technology

(CDT) to produce a live simulation for all attendees -- turning them into a trust & safety team,

tasked with responding to "reported" content on a fictional social media platform. Emma Llanso

from CDT and I ran the hour-long session, which included discussions of why people made their

eventual moderation decisions decisions. The video of that session can be viewed at

https://youtu.be/VIXGkoKfOS0.

52.   Many of the case studies we chose were designed to be challenging (most based on

real situations that content moderators had come across at various platforms). But the process was

useful and instructive. With each question there were four potential actions that the "trust &

safety" team could take and on every single example at least one person chose **each** option. In

other words, even when there was a pretty strong agreement on the course of action to take, there

was still at least some disagreement.

53.   Part of what we hoped to demonstrate was that when even a small group of

"professionals" couldn't even agree on just one out of eight cases, it demonstrated the

1 impossibility of content moderation at scale. As we noted, if you took that simulation and then

2 imagined (1) having to do that at scale, with hundreds, thousands, hundreds of thousands or even

3 millions of pieces of "flagged" content showing up, (2) having to do it when you're not someone

4 who is so interested in content moderation that you spent an entire day at a content moderation

5 summit, and (3) having to do it quickly where there are trade-offs and consequences to each

6 choice -- including possible legal liability -- and no matter which option you make, someone (or

7 perhaps lots of someones) is going to get very upset.

8   54.   This was not to say that internet platforms shouldn't strive to do better -- they

9 should. But one of the great things about attending these events is that it demonstrated how each

10 internet platform is experimenting in very different ways on how to tackle these problems. Google

11 and Facebook are trying to throw a combination of lots and lots of people plus artificial

12 intelligence at the problem. Wikipedia and Reddit are trying to leverage their own communities to

13 deal with these issues. Smaller platforms are taking different approaches. Some are much more

14 proactive, others are reactive. And out of all that experimentation, even if mistakes are being

15 made, we're finally starting to get some ideas on things that work for this community or that

16 community (and remember, not all communities work the same way).

17

18 **VII.   Platform Issues and Content Moderation**

19   55.   Large internet platforms like Redbubble need some level of moderation or they can

20 be overwhelmed with infringement, spam, or abuse. But at that scale, there will be a large number

21 of "mistakes" – both type I and type II errors (blocking content that shouldn't be blocked and

22 failing to block content that probably should be blocked) -- as well as disagreements over where to

23 come down over subjective issues regarding whether or not certain content should or should not be

24 allowed. While it will always be impossible for a platform like Redbubble to moderate in a way

25 that makes everyone happy, the company can still get *better* and make fewer errors. Part of the

26 way that companies learn to improve is through an iterative and experimental process of trying

27 different approaches to see which ones work best regarding the content on their own platform.

28

MASNICK EXPERT REPORT

1   56.   When it comes to allegations of intellectual property infringement, for example,

2 expecting a platform like Redbubble to identify all potential infringement on its own is impossible

3 -- which is one of the reasons why Congress established the takedown process of the DMCA,

4 requiring copyright holders to alert internet service providers of the specific instances of copyright

5 infringing material.   To identify whether a given listing on Redbubble's site infringes someone's

6 intellectual property, Redbubble first needs to know what that person claims their intellectual

7 property actually is. That's only the first step, of course, because the fact that someone *claims*

8 something is their intellectual property doesn't mean it is, and IP rightsholders often overstate the

9 scope of their rights, without regard to protecting legitimate criticism, parody, satire, or irony,

10 among other things.

11   57.   The fact that it's impossible to do content moderation at scale well **does not** mean

12 that companies like Redbubble shouldn't try to get better at it. They should. I have had the

13 opportunity to review Redbubble's policies and procedures, including its Atari-specific guidelines,

14 as well as filings from Redbubble's prior litigations against Ohio State University and a company

15 called LTTB, which describe both what Redbubble calls "proactive policing" – the company's

16 attempts to identify problematic content, including IP infringement, without being notified about it

17 by third parties – and its procedures for responding to takedown notices it receives from IP

18 rightsholders and others who complain about content hosted on the Redbubble site. I summarize

19 my understanding of those procedures as follows, in language adapted from Redbubble's filings in

20 the OSU and LTTB litigations.

21   58.   I understand that Redbubble's Marketplace Integrity ("MPI") Team proactively

22 polices the Redbubble Marketplace for potentially infringing content, using screening criteria

23 established by Redbubble. I understand that those screening criteria are based on information

24 Redbubble has received from content owners and almost always are created in collaboration with

25 those content owners. I understand that Redbubble performs proactive screening for Atari despite

26 Atari's refusal to cooperate in the process, making implementation of effective screening criteria

27 more difficult. Indeed, this level of proactive policing goes above and beyond the type of content

28 moderation efforts handled by other platforms.

MASNICK EXPERT REPORT

59.   I understand that Redbubble has created detailed internal policies and training for its MPI Team that teaches them how to spot signs of accounts that are more likely to post infringing content, and to disable such accounts quickly. I understand that these policies actually go above and beyond the policies of similarly situated companies.

60.   I understand that in Redbubble's proactive policing process, a member of the MPI Team creates a list of search terms that are deemed likely to correspond to potential infringing word marks or listings. I understand that this list of search terms is typically based on a list of claimed trademarks and other protected content provided by the content owner, and the content owner sometimes assists in providing common misspellings of its trademarks to make the search keywords more effective. I understand that if the content owner collaborates with Redbubble by providing it with a complete list of what it believes its protected content is, whether that be logos, trademarks, characters, etc., that means that Redbubble can search for and remove that content from the Marketplace more effectively. I understand that when content owners like Atari do not cooperate with Redbubble in this process, however, Redbubble is forced to use its best judgment as to what terms a content owner might use to identify arguable infringement. I have had the opportunity to review Redbubble's list of search terms for Atari, which appear to reflect a best efforts attempt to determine what types of content Atari is most worried about appearing on Redbubble.

61.   I understand that Redbubble personnel input the search terms into a proprietary software tool, which automatically and continuously monitors for changes in a dynamic database of Seller-generated titles, tags and descriptions in the Redbubble Marketplace, and in near real-time (i.e., approximately one minute from design upload to detection) displays the results of its searches in a user interface for the MPI Team to review and access. I understand that Redbubble's MPI Team manually reviews these search results, which include images of the artwork for the listing, to identify content that matches the policing guidelines, and is therefore potentially infringing, based on information provided by the content owner. I understand that, like the search terms, these guidelines are typically created in collaboration with the content owner, because policing functions most effectively if Redbubble has a full understanding of what designs a

MASNICK EXPERT REPORT

1   content owner considers infringing and how broadly the content owner wishes to enforce any

2   rights it has in such content, in part because various content owners have different approaches to

3   policing, particularly as it relates to fan art, mashups, and artwork that makes critical commentary

4   on the content owner's brand. I understand that if (as in the case of Atari) a content owner refuses

5   to work with Redbubble to create policing guidelines, Redbubble typically attempts to use its best

6   judgment as to what the content owner might consider arguably infringing and therefore might

7   want removed.

8   62.   I understand that if the policing tool displays user-generated content that matches

9   the content contained in policing guidelines, the MPI Team will promptly remove the listing

10   containing such content proactively and without receiving a specific takedown notice from the

11   rightsholder. I understand that in addition, the third-party Seller is automatically notified upon

12   removal and is otherwise treated in accordance with Redbubble's IP/Publicity Rights Policy.

13   63.   I understand that in 2017, Redbubble began testing a new internally developed tool

14   that allows a limited number of listings to be screened for potentially infringing titles or tags

15   immediately after upload but before they become publicly visible. I understand that this tool,

16   which Redbubble has sometimes referred to as the "ACE" tool, is still being tested and is under

17   review, and it has inherent limitations because it is keyword based and cannot match images. I

18   understand that, as a result, a human must still review each upload prior to removal to determine

19   whether the design matches any of the protected words or images in the policing guidelines

20   created in collaboration with the content owner. Given the high volume of designs uploaded daily

21   to the Marketplace – which I understand to have been nearly 19,000 listings on average in recent

22   years – it is my opinion that it would be logistically impossible for Redbubble employees to

23   perform a full pre-upload screening of **all** listings, and, in fact, doing so would likely increase the

24   error rate in decision making, by increasing the initial size of the data set on which "errors" would

25   be made.

26   64.   I understand that Redbubble also uses a combination of proprietary and third-party

27   software tools that seek to identify scaled and/or repeated abusers of the Redbubble user

28   agreement. For instance, I understand that Redbubble uses third-party machine learning software

1 called Sift Science to search for user accounts that are linked to other accounts that have already

2 been restricted or deleted for repeat infringement, as the existence of such accounts might indicate

3 that the applicable Sellers have attempted to circumvent Redbubble Policies by simply creating

4 new accounts. I understand that these tools also look for other Seller behavior that may indicate

5 that a Seller is a scaled abuser or repeat infringer, such as high content upload rates, which might

6 indicate that the user is not a legitimate human artist, but rather a software robot designed to

7 upload content in bulk. I understand that when Sift Science detects a high-risk account, it may

8 automatically disable the account or put it on a watch list to be monitored by the MPI Team. I

9 understand that when an account is disabled, all listings offered for sale by that account are

10 automatically removed from the Redbubble Marketplace.

11   65.   I understand that Redbubble has tested additional proprietary tools that can analyze

12 content, including image-based content (which is significantly more challenging than text-based

13 content), but has reasonably concluded that such tools are not currently adequate for content

14 moderation tasks. I understand that image-recognition tools are still a relatively new field in which

15 the only moderately successful products tend to rely on a large corpus of publicly accessible data

16 (e.g., Google's public image search), and that smaller proprietary attempts, focused on narrow data

17 sets (such as a collection of uploaded images for a single company) have yet to reach a level that

18 make them practical or effective.

19   66.   I understand that Redbubble's MPI Team currently conducts proactive policing for

20 marks belonging to Atari and approximately 200 other content owners, including some of the

21 largest content owners in the world. I understand that many of these companies have a large and

22 diverse portfolio of intellectual property rights (i.e., copyrights, trademarks and publicity rights) in

23 properties like TV shows and individual movies, and some represent a large roster of musical

24 artists. I understand that, in total, Redbubble's MPI team proactively polices for more than 1,500

25 individual properties for these content owners, and on any given day, the MPI Team may review

26 roughly 6,000 individual designs and compare them against the policing and removal guidelines

27 established (mostly) in collaboration with these content owners.

28

MASNICK EXPERT REPORT

67.   I understand that, as of April 2019, the proactive policing efforts of Redbubble's MPI Team have resulted in the disabling or removal of over 850,000 listings from the Redbubble Marketplace, covering more than 20 million products. I understand that, if listings in user accounts disabled by Sift Science are included, then proactive policing efforts as of that date had resulted in the disabling or removal of about 2,300,000 listings from the Redbubble Marketplace, covering approximately 66,700,000 products. And I understand that Redbubble had disabled and/or terminated roughly 468,000 Seller accounts for violation of Redbubble policies, including its IP/Publicity Rights Policy.

68.   I understand that, as of December 15, 2019, Redbubble has disabled or removed nearly 1000 listings on its marketplace – more than half of them under Redbubble's proactive removal guidelines – covering nearly 30,000 products, related to Atari's IP claims

69.   I understand that Redbubble frequently needs to make subjective decisions regarding what content should and should not be moderated. Quite frequently, the content at issue, even if it references Atari IP, might easily fall into a protected class of works, such as fair use, nominative fair use, or parody. This is true of both works that were proactively moderated by the Redbubble MPI Team, as well as moderation efforts in response to notices from Atari. I understand that Redbubble has educational material regarding such content, but that it is ultimately a subjective call to be made.

70.   As discussed earlier, with such content moderation decisions, there will always be some level of mistakes (Type I or Type II errors) made, and in reviewing a large list of decisions made by Redbubble, it appears that the company has leaned heavily towards being even more aggressive in proactively taking down content than might otherwise be necessary. That is, there are indications that Redbubble is even over-protecting and has chosen to have more Type I errors -- taking down content that might not need to be taken down -- in order to make sure that it is not running afoul of IP claims.

71.   Based on my review of these policies and procedures, and with the caveat that perfect moderation is impossible, it is my opinion that Redbubble's approach appropriately balances Atari's and other rightsholders' rights in its intellectual property and Redbubble's users'

1 rights as well. Indeed, it goes beyond what other, similarly situated companies do in attempting to

2 police infringing content. Content moderation involves choices, and every choice involves real

3 tradeoffs, and those tradeoffs can be significant and will upset some contingent who will have

4 legitimate complaints. From the outside, many people believe that content moderation involves

5 little more than having a single person or a small team dedicated to reviewing content -- and that

6 increasing the number of people can "solve" the problems associated with content moderation.

7 However, a site the size of Redbubble can't rely on a single person to review all the content that

8 needs review, and no single person can even know everything to look for. I believe that it is not

9 helpful to demand that companies magically do something that is impossible, especially when

10 much of that is trying to stop bad actors from working to try to game available systems to their

11 own advantage. No system can magically stop all bad actors. A serious look at the issues of bad

12 actors online starts with those actors and what they're doing, not in blaming the tools they use –

13 especially when those tools are content-neutral, such as with a print-on-demand service like

14 Redbubble.

15

16 **VIII.   Some Examples of the Impossibility of Content Moderation at Scale**

17   72.   Content moderation at scale is literally impossible to do well. It's not "difficult." It's

18 impossible. Although Redbubble is faced with an impossible task, it appears to do it as well as it is

19 reasonable to expect it to do. Indeed, it appears to go above and beyond what similar platforms of

20 similar sizes have done, and appears to take the impossible task seriously and approach it with an

21 eye towards always experimenting and improving.

22

23   *A.   Breasts on Facebook: The Quintessential Example of the Impossibility of*
           Content Moderation at Scale
24

25   73.   As I pointed out in a May 6, 2019 article on Techdirt, which I incorporate by

26 reference and adapt here, (https://www.techdirt.com/articles/20190503/17322942136/content-

27 moderation-scale-is-impossible-facebook-still-cant-figure-out-how-to-deal-with-naked-

28 breasts.shtml), and as noted earlier in this report, for over a decade, the quintessential example

1 used to show the impossibility of coming up with clear, reasonable rules for content moderation at

2 scale is Facebook's inability to figure out how to deal with pictures of naked breasts. In the early

3 days, as Facebook realized it needed to do some content moderation, and had to establish a clear

4 set of rules that could be applied consistently by a larger team, it started with a simple "no nudity"

5 policy -- and then after that raised questions, it was narrowed down to define female nipples as

6 forbidden.

7   74.   This might have seemed like a straightforward rule... until mothers posting

8 breastfeeding photos started complaining -- as they did after a bunch of their photos got blocked.

9 Stories about this go back at least until 2008 when the Guardian reported on the issue, after a

10 bunch of mothers started protesting the company, leading Facebook to come up with this

11 incredibly awkward statement defending the practice

12 (https://www.theguardian.com/media/2008/dec/30/facebook-breastfeeding-ban December 30,

13 2008):

14   "Photos containing a fully exposed breast, as defined by showing the nipple or
         areola, do violate those terms (on obscene, pornographic or sexually explicit
15   material) and may be removed," he said in a statement. "The photos we act upon
         are almost exclusively brought to our attention by other users who complain."
16

17   75.   More public pressure, and more public protests, resulted in Facebook adjusting its

18 policy to allow breastfeeding, but photos still kept getting taken down, leading the company to

19 have to keep changing and clarifying its policy, such as in this statement from 2012

20 (https://www.zdnet.com/article/facebook-clarifies-breastfeeding-photo-policy/).

21   When it comes to uploaded photos on Facebook, the vast majority of breastfeeding
         photos comply with our Statement of Rights and Responsibilities, which closely
22   mirrors the policy that governs broadcast television, and which places limitations
         on nudity due to the presence of minors on our site. On some occasions,
23   breastfeeding photos contain nudity – for example an exposed breast that is not
         being used for feeding – and therefore violate our terms. When such photos are
24   reported to us and are found to violate our policies, the person who posted the
         photo is contacted, and the photos are removed. Our policies strive to fit the needs
25   of a diverse community while respecting everyone¹s interest in sharing content that
         is important to them, including experiences related to breastfeeding.
26

27

28

76.   But photos of babies sleeping after having breastfed were getting taken down because the baby's head was no longer blocking the nipple.

77.   In 2014, Facebook clarified its policies on nipples again (https://www.salon.com/2014/06/15/facebook_frees_the_nipple_finally_allows_breast_feeding_photos_partner/ June 15, 2014):

> "Our goal has always been to strike an appropriate balance between the interests of people who want to express themselves with the interests of others who may not want to see certain kinds of content," a Facebook spokesperson told the Daily Dot. "It is very hard to consistently make the right call on every photo that may or may not contain nudity that is reported to us, particularly when there are billions of photos and pieces of content being shared on Facebook every day, and that has sometimes resulted in content being removed mistakenly.

> "What we have done is modified the way we review reports of nudity to help us better examine the context of the photo or image," the spokesperson continued. "As a result of this, photos that show a nursing mothers' other breast will be allowed even if it is fully exposed, as will mastectomy photos showing a fully exposed other breast."

78.   Then, just a few months later, people started protesting again, as, according to the Washington Post, more breastfeeding photos were taken down (https://www.washingtonpost.com/news/the-intersect/wp/2015/02/26/facebook-is-embroiled-in-yet-another-breastfeeding-photo-controversy/?utm_term=.adcbbc2644eb February 26, 2015).

79.   Then some people started posting photos of "breast feeding porn" that appeared to show breast feeding that wasn't infants. So they modified the rule to say the breastfeeding individual had to be an infant. But how does Facebook determine who is and who is not an infant? We're right back to the definitional problem. The original rule Facebook put in place was "does the kid look old enough to walk?" which raises other problems, since many kids breastfeed long after they can walk. Facebook has to keep amending and changing. It eventually allowed one (just one) nipple/areola showing if it appears related to breastfeeding. Later, Facebook changed the policy and said that a second nipple/areola could be shown.

80.   Facebook continues to grapple with this issue. Earlier in 2019 there were reports in Australia of some Facebook users were angry that Facebook was taking down a series of ads for

breast cancer survivors (https://www.9news.com.au/national/news-australia-facebook-breast-cancer-network-ads-ban-outrage-christchurch-terror-attack-video/c6f5d8f2-652c-43ea-b711-aa471f1d1649) . It appears that, over time, Facebook's rule has evolved so that there's some sort of amendment saying that there needs to be an educational component if you're showing breasts related to breast cancer.

81.   The charity in question called the whole thing "nonsensical," but it's actually the opposite of that. It's totally "sensical" once you understand much of the history, and the fact that Facebook keeps having to change and adapt these rules, often multiple times a month, to deal with the "new" cases that keep showing up that don't quite match. And while it seems "obvious" why these ads should be allowed, the company can't just rely on something being "obvious." It has over 10,000 people it employs who are in charge of making these decisions, and what's obvious to one of them may not be obvious to another. And thus it needs clearly spelled out rules. "Obvious" is not a workable rule.

82.   And those rules will **never** encompass every possible situation, and so these constant subjective calls and adjustments to the rules will continue. Content moderation at scale is impossible to do well, and part of that is because of stories like this. You can't create rules that work in every case, and there are more edge cases than you can possibly imagine.

B.   Other Examples

83.   Even when content moderation is done relatively well – as when Facebook dealt with the shooting in Christchurch, New Zealand – it's done imperfectly. (https://www.techdirt.com/articles/20190425/15315242087/behind-scenes-look-how-facebook-dealt-with-christchurch-shooting-demonstrates-impossible-task-content-moderation.shtml, Apr. 26, 2019).

84.   The French Internet Referral Unit used the Europol system to claim that much of the Internet Archive's site (which backs up and stores websites as a library-like archive) as "terrorist content" that violated the law and needed to be taken down. This included the Internet Archive's collection of CSPAN recordings, its Project Gutenberg collection of public domain

MASNICK EXPERT REPORT

1   books, the famous Prelinger Archive of public domain videos

2   (https://www.techdirt.com/articles/20190410/14580641973/eu-tells-internet-archive-that-much-

3   site-is-terrorist-content.shtml April 11, 2019)

4   85.   An Arizona candidate for the US Senate, Craig Brittain, who had settled a law

5   enforcement action against him by the FTC in 2015 over his running of a "revenge porn" site

6   called "Is Anybody Down?" (and proceeded to use a DMCA copyright claim to try to hide the

7   FTC's own report of the settlement from Google search results:

8   https://www.techdirt.com/articles/20150223/21334730120/deposed-revenge-porn-idiot-craig-

9   brittain-tries-to-censor-popehat-adam-steinbaugh-ftc.shtml Feb 24, 2015) also issued DMCA

10   notices seeking to take down embarrassing YouTube interviews of himself just days after he filed

11   to run for the Arizona Senate. (https://www.techdirt.com/articles/20170929/17100738318/former-

12   revenge-porn-site-operator-readies-senate-run-issuing-bogus-takedown-requests-to-youtube.shtml

13   Oct. 4, 2017)

14   86.   Law professor Eugene Volokh had to deal with multiple takedown notices trying to

15   hide his reporting and exposure of a "reputation management" company that tried to make use of

16   misleading takedown messages and even forged court documents to get content taken down.

17   (https://www.techdirt.com/articles/20190124/11043041460/guy-who-forged-court-order-to-delist-

18   content-issues-more-bogus-takedown-notices-to-remove-posts-discussing-his-forgery.shtml Jan

19   28, 2019)

20   87.   Earlier this year, Google sued and then settled with an individual who used threats

21   of false DMCA notices to extort YouTube users to pay him money.

22   (https://www.techdirt.com/articles/20191015/23132743201/guy-who-tried-to-extort-youtubers-

23   with-bogus-dmca-takedowns-agrees-to-settlement.shtml Oct. 17, 2019)

24   88.   The Electronic Frontier Foundation has created a website called "TOSsed Out" –

25   "TOS" being an acronym for "Terms of Service" – at https://www.eff.org/tossedout to track

26   stories of bad content moderation practices.

27

28

1   (https://www.techdirt.com/articles/20190521/23105242257/eff-highlights-stories-bad-content-

2   moderation-with-new-tossed-out-site.shtml May 22, 2019).

3

4   Executed this 20th day of December, 2019.

5   _Michael Masnick_____

6   Michael Masnick

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

Exhibit A

Gmail   **Leigh Beadon <leigh.beadon@gmail.com>**

Your Teespring campaign has been disabled

Leigh Beadon <leigh.beadon@floor64.com>   Fri, Dec 6, 2019 at 3:33 PM
To: ipescalations@teespring.com

Hi, I'm one of the operators of the Teespring account for rtb@techdirt.com

We received this notice (attached below) that our campaign at http://teespring.com/copying-is-not-theft was disabled for third-party content. This is an error - the campaign uses only original artwork that we created, and it has been live on Teespring for a long time now. No notice was attached so I am contacting you here as per instructions.

Please let me know when you can reinstate the campaign!

Thanks,
Leigh Beadon
Floor64/Techdirt
leigh.beadon@floor64.com

---------- Forwarded message ---------
From: **Teespring** <notifications@teespring.com>
Date: Thu, Dec 5, 2019 at 7:13 PM
Subject: Your Teespring campaign has been disabled
To: <rtb@techdirt.com>

Hello,

We are writing to you today because your campaign, Copying Is Not Theft, By Techdirt at http://teespring.com /copying-is-not-theft, has been terminated early due to content concerns. It appears your campaign may be using content owned by a third party.

We appreciate that intellectual property problems can happen inadvertently and have published some guidelines to help clarify how intellectual property works here http://answers.teespring.com/customer/en/portal/topics/936178-understanding-our-policies/articles – please take the time to familiarize yourself with these guidelines.

If we have included a notice attached to this notification and you disagree with the claimant and believe that the removal of the content is the result of a mistake (for example, that you have authorization) or misidentification, you can send us a counter notice by filling out a Counter-Notification Form https://teespring.com/policies/counter /claims/new. You can find information about the counter-notification process, the Counter-Notification form, and instructions for filling it out here http://answers.teespring.com/customer/en/portal/articles/2395971-notice-and-takedown-procedure. Failure to include information may result in rejection of your counter notice.

If we have not included a notice, and you believe the campaign was removed as a result of a mistake or you can show you have the right to reproduce the content in your campaign (for example, you have a written license from the third party rights holder), please let us know by emailing us at ipescalations@teespring.com as soon as possible. In that case, we may be able to reinstate your campaign. If you are not able to show you have the right to reproduce the content, we may not be able to reinstate your campaign.

We will fulfill any campaign orders placed prior to termination of the campaign. Any payment due to you will be shown in your payment dashboard.

Repeat violation of our policies will result in account termination, so please be sure to check any remaining campaigns for possible policy violations.

Thank you for your cooperation and understanding,

The Teespring Team

M Gmail   **Leigh Beadon <leigh.beadon@gmail.com>**

Your Teespring campaign has been disabled

IP Escalations <ipescalations@teespring.com>   Fri, Dec 6, 2019 at 6:02 PM
To: Leigh Beadon <leigh.beadon@gmail.com>

Hi,

You've been advised three times that the content has violated our acceptable use policy. You have been provided with the links to this policy for further information. This policy and choice to remove the content are not up for discussion. We apologize if you disagree with the decision. You will not receive anymore communication from us on this matter.

Thank you,
Team Teespring

On Fri, Dec 6, 2019 at 6:00 PM Leigh Beadon <leigh.beadon@gmail.com> wrote:
> If the "IP Escalations" team is unable to provide any further information, I would like to know where I escalate this matter to next.
>
> On Fri, Dec 6, 2019 at 5:38 PM Leigh Beadon <leigh.beadon@gmail.com> wrote:
>> I am not seeking a "debate" - I simply want to know what policy we violated. Please provide clarification.
>>
>> I am raising this issue with Teespring Support as well.
>>
>> On Fri, Dec 6, 2019 at 5:36 PM IP Escalations <ipescalations@teespring.com> wrote:
>>> Hi,
>>>
>>> We apologize if you disagree with our decision and for any inconvenience this matter has caused. Please understand that we are not in a position to debate our policies or discuss this issue further; however, your feedback has been noted and we truly appreciate your time today.
>>>
>>> Sincerely,
>>> Team Teespring
>>>
>>>
>>> On Fri, Dec 6, 2019 at 5:25 PM Leigh Beadon <leigh.beadon@gmail.com> wrote:
>>>> I have read the policies. The product does not violate any of them. Can you explain what violation you believe has occurred?
>>>>
>>>> On Fri, Dec 6, 2019 at 5:14 PM IP Escalations <ipescalations@teespring.com> wrote:
>>>>> Hi,
>>>>>
>>>>> You can find information regarding our our Acceptable Use Policies at https://teespring.com/policies/acceptable-use..
>>>>>
>>>>> Thank you,
>>>>> Team Teespring
>>>>>
>>>>> On Fri, Dec 6, 2019 at 4:29 PM Leigh Beadon <leigh.beadon@gmail.com> wrote:
>>>>>> Hi,
>>>>>>
>>>>>> We require some clarification. The email we received said the campaign contains third-party content - it does not. It is not in violation of any other acceptable use policy that I can see, and that was not what the notification email said anyway.
>>>>>>
>>>>>> Can you explain what policy this product violates?

-Leigh

On Fri, Dec 6, 2019 at 4:02 PM IP Escalations <ipescalations@teespring.com> wrote:
Hello Leigh,

Thank you so much for reaching out to Teespring today.

Unfortunately, we cannot reinstate the listing at this time.

 It appears your campaign may be in violation of our Acceptable Use Policies at https://teespring.com/policies/acceptable-use.

Please understand that in the interest of maintaining a safe and respectful platform, Teespring must take action quickly in response to complaints that a campaign may be in violation of our Acceptable Use Policies.

Repeat violation of our policies will result in account termination, so please be sure to check any remaining campaigns for possible policy violations.

Thank you for your cooperation and understanding.

The Teespring Team

On Fri, Dec 6, 2019 at 3:33 PM IP Escalations <ipescalations@teespring.com> wrote:
Hello,

Thank you for reaching out! We have received your request, and our Trust & Safety team is investigating your inquiry.

If you have not included the email address associated with your Teespring account as well as the name of the campaign in question, please reply to your original email and do so. We are unable to respond to emails that are missing such information.

If you have questions about why your campaign was removed, please see our Help Center article here.

Thank you,
Team Teespring

CERTIFICATE OF SERVICE

Atari Interactive, Inc. v. Redbubble, Inc.
United States District Court for the Northern District of California

STATE OF CALIFORNIA, COUNTY OF SAN MATEO

At the time of service, I was over 18 years of age and not a party to this action. I am employed in the County of San Mateo, State of California. My business address is 2000 Broadway St., Suite 154, Redwood City, CA 94063.

On December 20, 2019, I served true copies of the following document(s) described as **EXPERT REPORT OF MICHAEL MASNICK** on the interested parties in this action as follows:

Keith J. Wesley
Eric C. Lauritsen
Milin Chun
Matthew A. Venesia
BROWNE GEORGE ROSS LLP
2121 Avenue of the Stars, Suite 2800
Los Angeles, California 90067
	Counsel for Plaintiff Atari Interactive, Inc.
	
	Telephone: (310) 274-7100
	Facsimile: (310) 275-5697
	
	Email: kwesley@bgrfirm.com
	      elauritsen@bgrfirm.com
	      mchun@bgrfirm.com
	      mvenezia@bgrfirm.com

Kenneth B. Wilson COASTSIDE LEGAL 455 1st Avenue Half Moon Bay, CA 94019	*Counsel for Defendant Redbubble Inc.* Telephone: (650) 440-4211 Email: ken@coastsidelegal.com

BY E-MAIL OR ELECTRONIC TRANSMISSION: Based on a court order or an agreement of the parties to accept service by e-mail or electronic transmission, I caused the document(s) to be sent from e-mail address jmasur@zuberlawler.com to the persons at the e-mail addresses listed in the Service List. I did not receive, within a reasonable time after the transmission, any electronic message or other indication that the transmission was unsuccessful.

I declare under penalty of perjury under the laws of the United States of America that the foregoing is true and correct and that I am a member of the bar of this Court.

Executed on December 20, 2019, at Redwood City, California.

Joshua M. Masur